Intrusion Learning: An Overview of an Emergent Discipline

Tony Bailetti, Mahmoud Gad, and Ahmed Shah

> *"The illiterate of the 21st Century are not those who cannot read and write but those who cannot learn, unlearn and relearn.*
>
> Alvin Toffler
> Writer and futurist
> In *Powershift*

The purpose of this article is to provide a definition of intrusion learning, identify its distinctive aspects, and provide recommendations for advancing intrusion learning as a practice domain. The authors define intrusion learning as the collection of online network algorithms that learn from and monitor streaming network data resulting in effective intrusion-detection methods for enabling the security and resiliency of enterprise systems. The network algorithms build on advances in cyber-defensive and cyber-offensive capabilities. Intrusion learning is an emerging domain that draws from machine learning, intrusion detection, and streaming network data. Intrusion learning offers to significantly enhance enterprise security and resiliency through augmented perimeter defense and may mitigate increasing threats facing enterprise perimeter protection. The article will be of interest to researchers, sponsors, and entrepreneurs interested in enhancing enterprise security and resiliency.

Introduction

Intrusion learning offers the potential of significantly improving the security and resiliency of enterprise systems and increase the enterprise's capability to adapt to adversaries and changes in business environments. This article positions the emerging domain of intrusion learning at the intersection of machine learning, intrusion detection, and streaming network data. Machine learning refers to the algorithms that are first trained with reference input to "learn" its specifics, to then be deployed on previously unseen input for the actual detection process (Sommer & Paxson, 2010). Intrusion detection is the process of monitoring the events occurring in a computer system or network and analyzing them for signs of possible incidents, which are violations or imminent threats of violation of computer security policies, acceptable use policies, or standard security practices (Scarfone & Mell, 2007). By streaming network data, we mean streams of distinct and diverse network events flowing on a network over time. This definition is consistent with the definition of data stream provided by Savvius (2016).

We draw upon the results of a literature review carried out for the purpose of defining intrusion learning. We start with a summary of the literature review and then define intrusion learning, identify its distinctive aspects, and provide recommendations for advancing the emerging discipline. We end with our conclusions.

Literature Review

We performed a systematic narrative review to identify the latest advancements published in the academic literature with respect to machine learning, streaming network data, and intrusion detection. Articles in English-language journals published from 2010 to 2015 in North America and Europe were reviewed. We organized the literature into five themes: i) feature extraction, ii) learning algorithms, iii) clustering, iv) datasets, and v) tools.

Intrusion Learning: An Overview of an Emergent Discipline

Tony Bailetti, Mahmoud Gad, and Ahmed Shah

Feature extraction
Feature extraction is the process of determining a subset of features from an original set. The intent of feature extraction is to find a combination of original features or data attributes that can better describe the internal structure of the data. The three principal algorithms that are used for feature extraction are: locality preserving projection (linear projective maps arising from solving a variational problem optimally preserving neighbourhood structure), linear discriminate analysis (a method for finding a linear combination of variables that optimally separates classes) and principle component analysis (a linear technique that projects the data along the directions of maximal variance) (Fisher, 1936; He, 2005; Parakash & Surendran, 2013).

Intrusion detection systems use feature extraction to determine what features or attributes can assist with detecting malicious traffic (Laxhammer, 2014). We found two feature extraction challenges in the context of streaming network data. First, the dynamic changing nature of the streams results in challenges pertaining to the evolution of features (the emergence of new features), concept evolution (new classes evolving into the stream), and concept drift (underlying concepts change) (Momin & Hambir, 2015). The second challenge is that data streams are, in principle, of infinite length (Masud et al., 2010). Most existing data stream classification techniques address only the infinite length and concept-drift problems; concept evolution and feature evolution are ignored. In the face of a dynamic adversary, ignoring concept evolution and feature evolution increases enterprise risk.

Learning algorithms
Three emerging machine-learning algorithms play important roles in intrusion learning: active learning, adversarial learning, and conformal prediction. Active learning is a subfield of artificial intelligence and machine learning, and it refers to the study of computer systems that improve with experience and training (Settles, 2012). Adversarial learning refers to the study of effective machine learning techniques against an adversarial opponent (Huang et al., 2011). Conformal prediction refers to hedging individual predictions made by machine learning algorithms with valid measures of confidence (Laxhammar & Falkman, 2011).

The presence of an adversary changes the dynamics for learning algorithms. An adversary will attempt to poison or manipulate the data so that the algorithms treat the malicious as benign. This adversarial context has led to research on how algorithms can unlearn poisoned and polluted data (Cao & Yang, 2015).

Clustering
Organizing data into sensible groupings is one of the most fundamental modes of understanding and learning (Jain, 2010). Clustering is used to detect unknown attacks and discover unusual activities or usage patterns in traffic data in real time. The value of clustering comes from discovering groups and structures in the data that, in some way, are similar to each other, without prior knowledge of the data structures.

Data stream algorithms can only read the incoming data once and must do so in the context of having to respond in real-time with bounded memory usage. These algorithms can only provide approximate results and must support evolving concepts (Nguyen & Luo, 2013).

Because real-time data streams are unbounded, it will only be possible to process a portion of the entire data stream one "window" at a time (Nguyen & Luo, 2013). Various kinds of windows-based algorithms exist. For example, the sliding window algorithm analyzes the most recent data points and is suitable for applications where only the most recent information is of interest. The main disadvantage is that it ignores parts of the data streams. An adversary could manipulate a sliding window so that malicious activities occur in those parts of the streams being ignored by the algorithm.

Datasets
A dataset contains network traffic that is used to benchmark the performance of network intrusion algorithms. Datasets may include a combination of malicious traffic, non-malicious traffic, and identified features that can be used for testing. The most commonly used dataset researchers use for intrusion detection dates back to the KDD Cup 1999 (archive.ics.uci.edu/ml/datasets.html). It is surprising that a dataset from 1999 is still commonly used given the significant changes in attack tools, techniques, and data types that have occurred since then.

That the KDD Cup 1999 dataset is still used suggests that developing or accessing contemporary datasets is a major challenge. Privacy rights, confidentiality, and intellectual property are all concerns that impede access to real network data. Though there are other datasets available, the reality is that valid contemporary streaming data is unavailable outside of large Internet providers. The absence of new datasets retards science-based experimentation of new algorithms.

Intrusion Learning: An Overview of an Emergent Discipline

Tony Bailetti, Mahmoud Gad, and Ahmed Shah

Tools

Many publicly available experiments that are applying machine learning to intrusion detection are using a tool called massive online analysis (MOA; moa.cms.waikato.ac.nz). MOA is a machine-learning framework that contains real-time stream processing algorithms. It is not customizable for multi-node and scalable distributable processing.

However, scalable and distributable machine-learning processing engines that can process real-time streaming information do exist (e.g., SAMOA; samoa.incubator.apache.org). However, they have not been widely found in streaming intrusion-detection machine-learning experiments. We have not determined why this situation exists, though we note that SAMOA is a relatively new Apache project. SAMOA is one of few open source tools that is specifically designed for distributed and true real-time streaming (Landset et al., 2015). Apache Spark with MLib also includes a distributed architecture for processing data streams (spark.apache.org).

Defining Intrusion Learning

In this section, we propose a definition of intrusion learning based upon four elements: i) the ultimate outcome of intrusion learning; ii) the target of the ultimate outcome; iii) the mechanism used to deliver the ultimate outcome; and iv) the interdependence between intrusion learning and scientific and technological advances.

We propose the following definition of intrusion learning:

> *Intrusion learning is the collection of online network algorithms that learn from and monitor streaming network data resulting in effective intrusion detection methods for enabling the security and resiliency of enterprise systems. The network algorithms build on advances in cyber-defensive and cyber-offensive capabilities.*

We characterized the elements underpinning this definition as follows:

1. *Ultimate outcome:* Effective intrusion-detection methods on streaming network data.

2. *Target of ultimate outcome:* Security and resiliency of enterprise systems is the key target outcome.

3. *Mechanism used to deliver ultimate outcome:* Online network algorithms that learn from and monitor streaming network data.

4. *Interdependence of this mechanism from scientific and technological advances*: The mechanisms must build upon advances in both cyber-defensive and cyber-offensive capabilities (e.g., new machine-learning algorithms, new attack vectors), which themselves are informed by multi-disciplinary thinking.

Distinctive Aspects

We believe that there are five distinctive aspects of the intrusion learning domain relative to the machine learning, intrusion detection, and streaming domains:

1. *Real-time analysis of streaming network data:* Intrusion learning must respond to intrusions in real time. Unlike big data analytics, intrusion learning requires approximations, windowing, and other techniques to produce effective timely scalable analysis of network data (Aggarwal, 2007).

2. *High cost of failure:* The cost of failure of machine-learning algorithms is much higher for intrusion detection (e.g., loss of intellectual property and brand damage) compared to other applications of machine learning such as optical character recognition (Sommer & Paxson, 2010).

3. *Adversarial context:* Intrusion learning must deal with the existence of talented and determined adversaries. The presence of the adversary requires that intrusion learning must evolve with ongoing advances in both cyber-defensive and cyber-offensive capabilities (Cao & Yang, 2015; Corona et al., 2013).

4. *Network traffic diversity:* Intrusion learning must deal with the variability of network traffic (e.g., bandwidth, load balancing, and connection requests). Traffic diversity complicates the perspective of "normal" and therefore hinders the ability to identify an anomaly (Sommer & Paxson, 2010).

5. *Outlier detection:* Machine-learning algorithms are better at finding similarities than anomalies. As noted by Sommer and Paxon (2010), "the classic machine learning application is a classification problem, rather than discovering meaningful outliers as required by an anomaly detection system."

Intrusion Learning: An Overview of an Emergent Discipline

Tony Bailetti, Mahmoud Gad, and Ahmed Shah

Recommendations

The recommendations that follow are directed at researchers, sponsors and entrepreneurs interested in intrusion learning:

1. *Understand the threat model.* For example, researchers must know the cost of missed attacks (Sommer & Paxson, 2010).

2. *Learn, unlearn, and relearn.* Adversaries will act to mislead algorithms by steering the analyses to recognize the malicious as benign. Effective responses to such attacks need development. Corona and colleagues (2013) examine adversarial attacks against intrusion-detection systems as well as related taxonomies and potential solutions to known issues. This perspective leads to the concept of systems "unlearning" or forgetting what they had incorrectly "learned" (Cao & Yang, 2015).

3. *Select a narrow research scope.* The objectives of the research must be concrete. For example, researchers should determine precisely what kinds of attacks are being detected and what techniques are to be applied. The research should be able to answer such questions as to what attacks are being detected and the reasons as to why the attacks are being recognized (Sommer & Paxson, 2010).

4. *Develop new datasets.* To advance intrusion learning as a domain of practice, new datasets reflecting current network traffic need to be developed. For evidence-based evaluations, it is crucial to experiment with real datasets while observing societal norms such as privacy and commercial concerns.

5. *Develop open source intrusion learning tools that can scale.* Researchers need access to scalable machine learning tools. Although scalable proprietary tools exist, researchers worldwide must have access to tools that are capable of analyzing the reality of today's network traffic. Intrusion learning cannot advance in the absence of scalable machine learning tools.

6. *Improve online analytics.* Intrusion learning requires a combination of online and offline analyses. To properly enable real-time intrusion responsiveness, the balance between online and offline analytics needs to lean more heavily towards the online.

7. *Automate responses.* It is all very well to recognize the presence of anomalous or malicious activities. However, there is a need to go one step further and embed intrusion learning into the enterprise controllers. With highly scalable and changeable attacks, defensive responses must react in kind.

8. *Anticipate attacks.* By observing adversary community dynamics, it may be possible to anticipate attacks and react accordingly. Such research would move the discovery and detection outside the enterprise perimeter.

9. *Enhance feature extraction.* Research should aim to expand the set of extractable features that correlate with malicious traffic. This research could remain at the level of network flow, but richer theories are likely to provide more substantial payoffs.

Conclusion

In this article, we introduced the concept of intrusion learning as a domain that draws from machine learning, intrusion detection, and streaming network data. A key benefit of intrusion learning is that it may significantly enhance enterprise security and resiliency through augmented perimeter defense.

We identified a set of unique attributes and recommendations for advancing intrusion learning. For intrusion learning to meet its objectives of enhanced security and resiliency, these recommendations should not be treated in isolation but build upon each other: cross-cutting thinking (over machine learning, intrusion detection, and streaming) that focuses upon the distinctive aspects of intrusion learning will enhance progress.

Perhaps our most important recommendation is the development of new datasets that reflect contemporary network data and malware. The absence of such datasets is a significant impediment to the validation of intrusion-learning techniques. Privacy rights, confidentiality, etc., are concerns that are impeding the development of such datasets. We end this article with a "call to action" to develop such datasets, properly informed by researchers, privacy advocates, policy personnel, and so on, so that societal concerns are addressed.

Intrusion Learning: An Overview of an Emergent Discipline

Tony Bailetti, Mahmoud Gad, and Ahmed Shah

Acknowledgements

The authors thank Dan Craigen, Science Advisor at the Communications Security Establishment and a Visiting Scholar in the Technology Innovation Management program, for his invaluable input into the development and refinement of this article.

About the Authors

Tony Bailetti is an Associate Professor in the Sprott School of Business and the Department of Systems and Computer Engineering at Carleton University, Ottawa, Canada. Professor Bailetti is the Director of Carleton University's Technology Innovation Management (TIM) program. His research, teaching, and community contributions support technology entrepreneurship, regional economic development, and international co-innovation.

Mahmoud M. Gad is a Research Associate at VENUS Cybersecurity. He holds a PhD in Electrical and Computer Engineering from the University of Ottawa in Canada. Additionally, he holds an MSc in Electrical and Computer Engineering from the University of Maryland in College Park, United States. His research interests include cybercrime markets, machine learning for intrusion detection, analysis of large-scale networks, and cognitive radio networks.

Ahmed Shah holds a BEng in Software Engineering and is pursuing an MASc degree in Technology Innovation Management at Carleton University in Ottawa, Canada. Ahmed has experience working in cybersecurity research with the VENUS Cybersecurity Corporation and has experience managing legal deliverables at IBM.

References

Aggarwal, C. (Ed.) 2007. Data Streams: Models and Algorithms. New York: Springer.

Cao, Y., & Yang, J. 2015. Towards Making Systems Forget with Machine Unlearning. In *Proceedings of the 2015 IEEE Symposium on Security and Privacy:* 463–480. New York, IEEE. http://dx.doi.org/10.1109/SP.2015.35

Corona, I., Giacinto, G., & Roli, F. 2013. Adversarial Attacks Against Intrusion Detections Systems: Taxonomy, Solutions and Open Issues. *Information Science,* 239: 201–225. http://dx.doi.org/10.1016/j.ins.2013.03.022

Fisher, R. A. 1936. The Use of Multiple Measurements in Taxonomic Problems. Annals of Eugenics, 7(2): 179–188. http://dx.doi.org/10.1111/j.1469-1809.1936.tb02137.x

He, X. 2005. *Locality Preserving Projections.* Doctoral thesis, University of Chicago.

Huang, L., Joseph, A. D., Nelson, B., Rubinstein, B. I. P., & Tygar, J. D. 2011. Adversarial Machine Learning. In *Proceedings of the 4th ACM Workshop on Security and Artificial Intelligence:* 43–58. http://dx.doi.org/10.1145/2046684.2046692

Jain, A. 2010. Data Clustering: 50 Years Beyond K-means. *Pattern Recognition Letters,* 31(8): 651–666. http://dx.doi.org/10.1016/j.patrec.2009.09.011

Landset, S., Khoshgoftaar, T. M., Richter, A. N., & Hasanin, T. 2015. A Survey of Open Source Tools for Machine Learning with Big Data in the Hadoop Ecosystem. *Journal of Big Data,* 2(1): 1–36. http://dx.doi.org/10.1186/s40537-015-0032-1

Laxhammer, R. 2014. *Conformal Anomaly Detection: Detecting Abnormal Trajectories in Surveillance Applications.* Doctoral Thesis, University of Skövde School of Informatics, Sweden.

Laxhammar, R., & Falkman, G. 2011. Sequential Conformal Anomaly Detection in Trajectories Based on Hausdorff Distance. In *Proceedings of the 14th International Conference on Information Fusion.* New York, IEEE.

Masud, M., Chen, Q., Guo, J., Khan, L., & Han, J., Thuraisingham, B. M. 2010. Classification and Novel Class Detection of Data Streams in a Dynamic Feature Space. In *Proceedings of the European Conference on Machine Learning and Knowledge Discovery in Databases:* 337–352, 2010. http://dx.doi.org/10.1109/TKDE.2010.61

Momin, N., & Hambir, N. 2015. A Survey on Various Classification and Novel Class Detection Approaches for Feature Evolving Data Stream. *Multidisciplinary Journal of Research in Engineering and Technology,* 2(1): 342–346.

Nguyen, K., & Luo, Z. 2013. Reliable Indoor Location Prediction Using Conformal Prediction. *Annals of Mathematics and Artificial Intelligence,* 74(1): 133–153. http://dx.doi.org/10.1007/s10472-013-9384-4

Parakash, D., & Surendran, S. 2013. Detection and Analysis of Hidden Activities in Social Networks. *International Journal of Computer Applications,* 77(16): 34–38. http://dx.doi.org/10.5120/13570-1404

Savvius. 2016. Glossary of Networking Terms. *Savvius.* Accessed February 15, 2016: http://www.wildpackets.com/resources/compendium/glossary_of_networking_terms#S

Intrusion Learning: An Overview of an Emergent Discipline

Tony Bailetti, Mahmoud Gad, and Ahmed Shah

Scarfone, K., & Mell, P. 2007. *Guide to Intrusion Detection and Prevention Systems (IDPS)*. NIST Special Publication 800-94. Gaithersburg, MD: National Institute of Standards and Technology.

Settles, B. 2012. Active Learning. *Synthesis Lectures on Artificial Intelligence and Machine Learning,* 6(1): 1–114. http://dx.doi.org/10.2200/S00429ED1V01Y201207AIM018

Sommer, R., & Paxson, V. 2010. Outside the Closed World: On Using Machine Learning for Network Intrusion Detection. In *Proceedings of the 2010 IEEE Symposium on Security and Privacy:* 305–316. http://dx.doi.org/10.1109/SP.2010.25

Citation: Bailetti, T., Gad, M., & Shah, A. 2016. Intrusion Learning: An Overview of an Emergent Discipline. *Technology Innovation Management Review*, 6(2): 15–20. http://timreview.ca/article/964

Keywords: cybersecurity, intrusion learning, intrusion detection, machine learning, learning algorithms, adversarial learning, clustering, streaming network data, real-time analysis, enterprise, security, resiliency